Methodology for Determination of the Upper Safety Limit for Criticality Calculations for Criticality Safety Analyses

R. H. Smith
H. J. Keener
J. F. DeClue
A. W. Krass

Nuclear Criticality Safety Division
Y-12 National Security Complex

April 2001

DISCLAIMER

This work of authorship and those incorporated herein were prepared by Contractor as accounts of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, use made, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency or Contractor thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency or Contractor thereof.

COPYRIGHT NOTICE

This document has been authored by a subcontractor of the U. S. Government under contract DE-AC05-00OR-22800. Accordingly, the U. S. Government retains a paid-up, nonexclusive, irrevocable, worldwide license to publish or reproduce the published form of this contribution, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, or allow others to do so, for U. S. Government purposes.

**Methodology for Determination of the Upper Safety Limit for
Criticality Calculations for
Criticality Safety Analyses**

R. H. Smith
H. J. Keener
J. F. DeClue
A. W. Krass

Balance of Plant Group
Nuclear Criticality Safety Division
Y-12 National Security Complex

April 2001

THIS PAGE INTENTIONALLY LEFT BLANK

Document Preparation

Prepared by _____[signature on file]_____ Date _____
         Robert H. Smith

Reviewed by _____[signature on file]_____ Date _____
         John F. DeClue

Approved by _____[signature on file]_____ Date _____
         Young Y. Chan

THIS PAGE INTENTIONALLY LEFT BLANK

Abstract

This report considers the methods for determination of an upper safety limit, and incorporating uncertainty and margin into the safety limit, provides comparisons, and recommends a preferred method for determining the Upper Safety Limit (USL).

A USL is developed for CSAS25 from SCALE4.4a. The USL is applicable for the CSAS25 control module from the SCALE 4.4a computer code system for use in evaluating nuclear criticality safety of enriched uranium systems. The benchmark calculation results used for this report are documented in Y/DD-896. The statistical evaluation is documented in CCG-380. The 27-group ENDF/B-IV, 44-group ENDF/B-V, and 238-group ENDF/B-V cross-section libraries were used.

Numerical methods for applying margins are described, but the determination of appropriate correlating parameters and values for additional margin, applicable to a particular analysis, must be determined as part of a process analysis. As such, this document does not specify final upper subcritical limits as has been done in the past. No correlation between calculation results and neutron energy causing fission was found for the critical experiment results. Analysts using these results are responsible for exercising sound engineering judgment using strong technical arguments to develop "a margin in k_{eff} or other correlating parameter … that is sufficiently large to ensure that conditions (calculated by this method to be subcritical by this margin) will actually be subcritical."[1] Documentation of area of applicability and determination and justification of the appropriate margin in the analyst's evaluation, in conjunction with this report, will constitute the complete Validation Report in accordance with ANSI/ANS-8.1-1998, Section 4.3.6(4).

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

THIS PAGE INTENTIONALLY LEFT BLANK

1 Introduction

In the nuclear criticality safety field, extensive use is made of computational methods in order to establish subcritical limits for operations involving fissile material. Calculations must be made by a method shown to be <u>validated</u> by comparison with experiment data. The American National Standards Institute (ANSI) gives fundamental guidance for establishing these subcritical limits in ANSI/ANS-8.1-1998, ANuclear Criticality Safety in Operations with Fissionable Materials Outside Reactors.@[1]

The successful application of computational methods requires verification, certification and validation (V&V) of the analytic tool. Software verification, the first element of V&V, reveals the closeness in agreement between calculation results obtained using a particular computational method and calculation results obtained using other methods. Software certification provides a user assurance that the code will <u>consistently</u> produce numerical results that agree with analytic solutions for the problems the code is designed to analyze. Software validation Acalibrates@ the calculation results obtained using certified software (computational method) for agreement with the physical universe.

For nuclear criticality safety, the validation process is complete with development of the upper safety limit (USL) from validation data. Refinement of a validation or extension by addition of validation data does not require re-certification of the software. Therefore, the verification and certification of a nuclear criticality safety computer code system and transference of a validation between computer platforms are related issues, which can be addressed separately from software validation.

Taken together, Y/DD-896 and Y/DD-972 provide the **framework** for the validation of CSAS25 control module from the Standardized Computer Analyses for Licensing Evaluation (SCALE) Version 4.4a package with the 27-group ENDF/B-IV, 44-group ENDF/B-V, or 238-group ENDF/B-V SCALE 4.4a cross-section libraries.

Y/DD-896 addresses the establishment of bias, bias trends and uncertainty associated with the use of SCALE4.4a for performance of criticality calculations. This evaluation is directed at enriched uranium systems consisting of fissile and fissionable material in metallic, solution and other physical forms, as described in the Organization for Economic Cooperation and Development (OECD) Handbook. The focus is on comparison of calculated neutron multiplication factors (k_{eff}) with the associated experimental results, and on establishment of bias, bias trends, and uncertainty as a final step. Compiled data for 503 critical experiments is used as the basis for the calculation models. The calculated results from SCALE 4.4a using the 27-group ENDF/B-IV, 44-group ENDF/B-V, and 238-group ENDF/B-V SCALE 4.4a cross-section libraries have been compared with

reported results for the benchmark experiments. Comparison of results demonstrates that SCALE 4.4a run on the Nuclear Criticality Safety Division (NCSD) Hewlett Packard Series 9000/J-5600 (HP J-5600) unclassified workstation (referred to as CMODB) produces the same results within the statistical uncertainty of the Monte Carlo calculations as reported by the experiments.

Y/DD-972 addresses various methods for determination of a USL and for incorporating uncertainty and margin into this USL. It provides comparisons of these methods and recommends a preferred method for determining the USL. Y/DD-972 establishes subcritical limits determined through an evaluation of statistics parameters of calculation results for critical experiments. However, the correlating parameters (i.e., mass, enrichment, geometry, absorption, moderation, reflection, etc.) and values for applying additional margin to the subcritical limits are application dependent. The determination of correlating parameters and additional margin is an integral part of the process analysis for a particular application. For the critical experiment results, no correlation between calculation results and neutron energy causing fission was found. As such, this document does not specify Afinal@ USL values as has been done in the past.

The AArea of Applicability@ (AoA) is also application dependent. In using the "raw" USL data of this report, the nuclear criticality safety analyst must exercise sound engineering judgment and use strong technical arguments to develop a margin in k_{eff} or other correlating parameter.   That margin must be sufficiently large to ensure that conditions (calculated by this method to be subcritical by this margin) will actually be subcritical.[1] Preparation of the final portion of the "written report of validation" as identified in Section 4.3.6 of ANSI/ANS-8.1-1998 is the responsibility of the analyst.

1.1   Grandfather or Retrofit Concerns

The validation documents used for establishing subcritical limits for criticality calculations shall be reviewed and approved by NCSD. There is no urgency to retrofit older evaluations since the common practice provided conservatism. Therefore, the incorporation of these methods into the evaluations of existing operations will be accomplished as the Nuclear Criticality Safety Approvals are updated due to equipment changes, rule changes, age, etc.

1.2   Terminology

The term "model" is used in two different ways in this report. In the main body of the report, "model" generally refers to the calculation model, or the collection of input parameters for the numerical simulation of the physical critical experiment. In the statistical analysis in Appendix A, the term "model" generally refers to a statistical model, or an equation describing the relationship between parameters considered in the statistical analysis.

The term "uncertainty" is used two different ways in this report. In the main body of the report, "uncertainty" refers to less than exactly known values in the calculation model, or in the collection of input parameters for the numerical simulation of the physical critical experiment. In the statistical analysis in Appendix A, the term "uncertainty" generally refers to the variation of repeated observations (e.g.,k_eff, k_sim, k_plt) about their expected value. This uncertainty is quantified by the sample standard deviation (i.e., square root of the sample variance).

Accuracy refers to how close the expected value of an observation is to the true value. Accuracy is quantified by bias, which is the difference between the sample average and the true or referenced value.

2   Determination of Subcritical Limits and Discussion of Statistical Analysis

2.1   Introduction

Y/DD-896 provides a discussion of the calculation results, experiment (benchmark) data, and the bias results. Regression with neutron energy causing fission was found to be not applicable. The statistical analysis is documented in CCG-380, which is included with this report, as Appendix A. The first part of CCG-380 is also included as an appendix to Y/DD-896.

The OECD reports provide an estimate of the actual k_{eff} for each experiment. While most experiments were estimated to be exactly critical, or k_exp = 1.0, a number of the estimates were slightly above or slightly below exactly critical (i.e. k_exp > 1.0 or k_exp < 1.0). It was considered desirable to account for this difference because the subcritical limit could otherwise be conservative due to including slightly subcritical experiments, or non-conservative due to including slightly super-critical experiments. The following adjustment, to determine an adjusted calculation result, k=, was proposed for possible use in the statistical evaluation (Appendix A; or CCG-380):

k' = 1 + bias = 1 + (k_calc - k_exp)

where k_calc = code calculated k_{eff} result for the modeled experiment, and
        k_exp = OECD estimate of actual k_{eff} for the experiment

The assumption inherent in making this adjustment is that the experiments were very close to critical, and if a very small adjustment could be made so that the experiment would be exactly critical, the bias of the calculation would not change. In the statistical evaluation, the bias and uncertainty was determined.

When calculation methods are employed, the subcritical limit is a quantitative value used for

implementation of NCSD policy, which is established in accordance with DOE Orders, Standards, or by NRC regulations for those activities under NRC jurisdiction. Evaluation criteria are usually the calculated reactivity, k_{eff}, for the model of the problem being evaluated, uncertainty in the calculated value of reactivity, bias introduced by the code used to calculate reactivity, and bias introduced by the model. Code bias is determined by a validation performed in accordance with ANSI/ANS 8.1. The statistical results in this report determine code bias based on comparison of code calculations with experimental results. Model bias is due to uncertainties in the modeled parameters describing the problem, approximations made for the model to conform to the input requirements of the code, and any significant simplifications of the problem introduced into the model by the analyst. The acceptance criterion requires that the evaluation criteria be bounded by the subcritical limit. In the case of code calculations, the subcritical limit is often a criterion based on calculated k_{eff} and estimated bias and uncertainties.

It is recognized that it may also be possible to determine subcritical limits based on adjustment of correlating parameters other than k_{eff}, as allowed by ANSI/ANS-8.1. These parameters could include physical parameters such as mass or radius, for example, or other calculated parameters besides k_{eff}, if they can be shown to provide a subcritical margin. This report does not address other potential correlating parameters, but an analyst could determine them as part of an evaluation and use them to determine subcritical limits in conjunction with the code bias results in this report.

2.2   Determination of Subcritical Limits from the Bias and Uncertainty

Several methods could be used for determination of subcritical limits in terms of calculated k_{eff}. The methods will be examined in more detail, and an approach will be recommended in this report. As discussed in Section 5.1 of Y/DD-896, it may also be possible to determine subcritical limits based on adjustment of correlating parameters other than k_{eff}, but this approach is not addressed here.

2.2.1   Area of Applicability

Whether a particular plant application is similar to a sufficient number of experiments contained in the validation set is an important issue. This is a prerequisite before a subcritical limit may be determined from the validation results. Based on the demonstrated performance of the code for a large number of experiments, we have reasonable assurance that it performs well over a wide range of conditions. However, this does not ensure that the code performs as well over all conditions. It does not guarantee that there is not a scenario where the code would calculate unusually low due to some kind of anomaly or error in the code or cross section data. There are some types of experiments that are one or two of a kind. It is possible that there could be one or more of these types of experiments that would exhibit a larger bias than seen in this report, if more

experiments of that type were included. The only way to assess whether the validation set contains a sufficient number of experiments that are similar to a plant application is to carefully compare the details of the experiment description with the details of the plant application. The broad categories used for the statistical analysis are not sufficient for making the determination, and even the experiment summary information included in this report may not provide enough information. The analyst must become familiar with the experiment descriptions provided in the OECD handbook. The minimum number of similar experiments needed cannot be determined in advance, but the analyst must consider this issue carefully as one of the factors when deciding how much margin to include in determining a subcritical limit. If there are only a few experiments similar to the application, it may be wise to recommend taking the time to select additional applicable experiments for addition to the validation set, or for a separate evaluation.

Appropriateness of the subcritical limit does not depend solely on the validity of the statistical analysis. Rather, the appropriateness of a subcritical limit depends on two quite different factors: (1) the range of conditions under which a generalization for its use is established, and (2) whether or not the modeled scenario is within the range of these conditions. The nuclear criticality safety analyst exercising sound engineering judgment using strong technical arguments must develop these determinations. Extrapolation beyond the range of applicability should not be done.

2.2.2 Simple Methods

The simplest and most straightforward method is to observe that for a large number of critical (or close to critical) experiments, calculated results were relatively consistent, and none were lower than a particular minimum. Limits would be established based on the cross section library and physical form. A value of k_{eff} could be picked that equals or is slightly below all the calculated results. Combined limits could be used for simplicity as long as the minimum values were used. The results of this approach are shown in columns 3-5 of Table 1. The minimum value is established by a single case. These values do not include any reduction for numerical uncertainty in the calculated mean value for the case (standard deviation of approximately 0.001). The minimum k_{eff} value would be reduced further to provide additional margin as appropriate as determined by the analyst.

Another similar method to establish a subcritical limit would be to choose the least positive or most negative bias result (lowest bias), add it to 1.0, and add a further reduction to allow for a margin of safety. The results of this approach (without reduction for margin of safety) are shown in columns 6-8 of Table 1. These values do not include any reduction for numerical uncertainty in the calculated mean value for the minimum case.

2.2.3 Addition of Safety Margin and Bias Based on Tolerance Bounds

In the commonly used relationship of procedure Y70-68-005, the subcritical limit appears in Equation 2.1 as the term on the right while the evaluation criteria of a given problem are expressed in the term on the left. It is recognized that, depending on how the terms are defined, this approach may mix confidence intervals ($k_{eff} + 2\sigma$) and tolerance limits (1.0 - bias - margin) in a semi-statistical formula or method. Acceptability is established when the term on the left is less than or equal to the term on the right:

$$\{ k_{eff} + 2\sigma + \text{code bias} + \text{model bias} \} \leq \{ 1.0 - \text{margin of safety} \} \quad \text{(Eq 2.1)}$$

Code bias is determined by a validation performed in accordance with ANSI/ANS 8.1. The statistical results in this report (Appendix A) present bias based on comparison of code calculations with experimental results. Some regulators require that when the mean value of calculated k_{eff} for the validation cases is greater than or equal one, then the code bias is set = 0. See Appendix A for details on determination of the code bias.

Model bias is due to uncertainties in the modeled parameters describing the problem, approximations made for the model to conform to the input requirements of the code, and any significant simplifications of the problem introduced into the model by the analyst, as discussed in Section 2.2.3.1.

Margin of safety is set by NCSD analysts in accordance with DOE Orders, Standards, and NRC Regulations for activities under these specific jurisdictions. A typical range is from 0.02 to 0.05. When the analyst feels there is very little uncertainty in the application, it may be possible to justify using lower margin of safety. Also, it may be appropriate to provide margin by adjustment of correlating parameters other than k_{eff}.

2.2.3.1

For calculational models that are accurate representations, the model bias is considered zero. calculational models with uncertainties in the modeled parameters describing the problem, model bias is intended to compensate k_{eff}
dominant parameter(s) of the problem. Quantitatively, model bias is determined by evaluating k of problem "A" as a function of the parameter "p" over the range of uncertainty in "p" [expressed here %], i.e., calculations performed with the parameter "p" set at values between "p(1-x%/100)" and k_{eff}
extremes must be considered because reactivity may not be uniformly distributed about the parameter "p". The number of values to be evaluated depends on the magnitude of the range in the parameter.

difference:

$$| \ k_{eff} \ \{ \text{"p\_max"} \} - \ k_{eff} \ \{ \text{"p\_min"} \} \ | ,$$

Where $k_{eff}\{\text{"p\_max"}\}$ = maximum k_{eff} over the range of parameter(s) "p"
$k_{eff}\{\text{"p\_min"}\}$ = minimum k_{eff} over the range of parameter(s) "p"

The maximum variation in k_{eff} over the ranges of uncertainty of all dominant parameters of the problem would be used as the model bias in Equation 2.1. Allowance for statistical uncertainty in the calculated results must be considered, depending on the relative magnitude. If the parameter of interest is a discrete or categorical variable rather than a continuous variable as discussed above, then the parameter would similarly be varied over the range of applicable states until the maximum variation was found.

An allowance for model bias could otherwise be included in the margin of safety as shown in Eq. 2.2:

$$\{ k_{eff} + 2\acute{o} + \text{code bias} \} \ \leq \ \{ 1.0 - \text{margin of safety*} \} \qquad \text{(Eq 2.2)}$$

This would require justification that the evaluation criteria (left side of Eq. 2.2) is sufficiently less than 1.0 by an amount out weighing the maximum variation of k_{eff} for the dominant parameters affecting reactivity.

For calculational models with a significant degree of approximation and simplification, model bias is intended to account for conservatism built into the model. Generally, the model bias is not explicitly calculated, but is addressed by modeling parameters at their worst case. Establishment of the worst case may require a parametric study as described previously. An allowance for model bias would be included in the margin of safety (term on right side of Eq. 2.2). This would require justification that the evaluation criteria (left side of Eq. 2.2) is sufficiently less than 1.0 by an amount that out weighs the conservatism built into the calculational model.

2.2.4 Statistical Approach

CCG-380 formulates a "Statistical Model" and incorporates uncertainties in a rigorous way. It is included as Appendix A to this report, and selected results are incorporated in the following discussions.

2.2.5 Discussion of Results

Based on the methods discussed previously, Table 1 summarizes the results that are relevant for choosing an USL before allowance for additional margin. The first two columns show the enrichment

and physical form category, and the number of simulations included in the category. Columns 3-5 show the minimum calculated result with each of the cross section libraries. Columns 6-8 shows the minimum value of the quantity (1+bias), where bias = k_cal - k_exp. Columns 9-11 shows the statistics based Lower Tolerance Limit (LTL) for at least 0.95/95%, or USL without any additional allowance for safety margin.

Although the results for the 27-group library are presented and discussed for comparison purposes in this report, as discussed in Y/DD-896 it is recommended that analysts begin using the 44- or 238-group libraries. In general, the 238-group is preferred unless computer system performance necessitates a compromise, or for applications where the 44-group library may perform better (Low Enriched LWR and MOX lattices, and ^{233}U.)

Although four decimal places are shown in Table 1, the actual accuracy of the code, for a particular calculation, may be on the order of +/- 0.02 to 0.03 based on the spread in the bias results. Also, the standard deviation of the mean for a particular calculation was on the order of 0.001 for benchmark cases, and somewhat higher for applications. Therefore, less than four decimal places are physically meaningful or significant for developing limits. Too much importance should not be placed on small differences in the third or fourth decimal place. The main value to reporting four (or five) decimal places for a calculation result is to confirm that the result reported actually originated from a given output file.

Comparing the results in Table 1 columns 3-5 versus columns 6-8, most corresponding results are close to the same values. The values that are "significantly" different are highlighted in columns 6-8. Essentially, this involves only the HEU Metal experiments. In context here, a difference of less than about 0.0012 is considered not significant. The HEU Compound 27-group values differ by only 0.0012, and several of the LEU values differ by less than 0.0005. Accounting for the experiment k-effective being different from 1.0 "significantly" changes only the HEU Metal results. This change also effects the Combined Category for all HEU and LEU experiments.

Comparing the results in Table 1 columns 6-8 versus columns 9-11, most corresponding values are again close to the same values. Most corresponding values are identical. The values that are not identical are again highlighted in columns 9-11. These differences occur for the several categories that were normally distributed according to the statistical evaluation (CCG-380 and Appendix A).

The finding of normal distribution is considered to be interesting for the following reasons. Six of the seven normally distributed categories had very few experiments, either 22 or 9. The seventh normally distributed category, HEU Solutions with the 27-group library, had a large number of experiments, 267. Unlike the other six categories, in the case of the HEU solutions with the 27-group library the statistics based LTL is actually higher than the minimum 1+bias value (0.9947 versus 0.9844). Therefore, for the 27-group HEU solutions it is conservative to use the minimum bias (1+bias) value.

Table 1. Summary of Calculated k-effective Results for Upper Safety Limit Determination, Before Allowance for Margin

Physical Form By Uranium Enrichment	Number of Simulations	27-group Minimum Calculated Result	44-group Minimum Calculated Result	238-group Minimum Calculated Result	27-group Minimum 1+bias Result	44-group Minimum 1+bias Result	238-group Minimum 1+bias Result	27-group Statistics Based LTL (at least 0.95/95%)	44-group Statistics Based LTL (at least 0.95/95%)	238-group Statistics Based LTL (at least 0.95/95%)
High Enriched Uranium Metal	118	0.9780	0.9717	0.9681	0.9908	0.9846	0.9797	0.9908	0.9846	0.9797
Compound	22	0.9874	0.9903	0.9858	0.9886	0.9903	0.9858	0.9771	0.9895	0.9845
Solution	267	0.9844	0.9812	0.9802	0.9844	0.9812	0.9802	0.9947	0.9812	0.9802
Combined Category for High Enriched Uranium	407	0.9780	0.9717	0.9681	0.9844	0.9812	0.9797	0.9844	0.9812	0.9797
Low Enriched Uranium Metal	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Compound	87	0.9769	0.9804	0.9751	0.9769	0.9804	0.9753	0.9769	0.9804	0.9753
Solution	9	0.9934	0.9944	0.9941	0.9929	0.9939	0.9936	0.9894	0.9902	0.9894
Combined Category for Low Enriched Uranium	96	0.9769	0.9804	0.9751	0.9769	0.9804	0.9753	0.9769	0.9804	0.9753
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	0.9769	0.9717	0.9681	0.9769	0.9804	0.9753	0.9769	0.9804	0.9753

Notes: (1) Shaded values in columns 6-8 indicate values that are very different from corresponding values in columns 3-5.
(2) Shaded values in columns 9-11 indicate values based on normal distributions.

Table 2. Potential Upper Subcritical Limits, Before Allowance for Margin

Physical Form By Uranium Enrichment	Number of Simulations	27-group	44-group	238-group
High Enriched Uranium				
Metal	118	0.990	0.984	0.979
Compound	22	0.977	0.989	0.984
Solution	267	0.984	0.981	0.980
Combined Category for High Enriched Uranium	407	0.984	0.981	0.979
Low Enriched Uranium				
Metal	N/A	N/A	N/A	N/A
Compound	87	0.976	0.980	0.975
Solution	9	0.989	0.990	0.989
Combined Category for Low Enriched Uranium	96	0.976	0.980	0.975
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	0.976	0.980	0.975

Table 3. Recommended USL Example, Assuming 0.02 Additional Margin of Safety

Physical Form By Uranium Enrichment	Number of Simulations	27-group	44-group	238-group
High Enriched Uranium Metal	118	0.970	0.964	0.959
Compound	22	0.957	0.969	0.964
Solution	267	0.964	0.961	0.960
Combined Category for High Enriched Uranium	407	0.964	0.961	0.959
Low Enriched Uranium Metal	N/A	N/A	N/A	N/A
Compound	87	0.956	0.960	0.955
Solution	9	0.969	0.970	0.969
Combined Category for Low Enriched Uranium	96	0.956	0.960	0.955
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	0.956	0.960	0.955

3 Conclusions

Taken together, Y/DD-896 and Y/DD-972 provide the **framework** for the validation of CSAS25 control module from the Standardized Computer Analyses for Licensing Evaluation (SCALE) Version 4.4a package with the 27-group ENDF/B-IV, 44-group ENDF/B-V, or 238-group ENDF/B-V SCALE 4.4a cross-section libraries. (Similar documentation will be developed for MCNP.)

Y/DD-896 addresses the establishment of bias, bias trends and uncertainty associated with the use of SCALE4.4a for performance of criticality calculations. This evaluation is directed at enriched uranium systems consisting of fissile and fissionable material in metallic, solution and other physical forms, as described in the OECD Handbook. The focus is on comparison of calculated neutron multiplication factors (k_{eff}) with the associated experimental results, and on establishment of bias, bias trends, and uncertainty as a final step. Compiled data for 503 critical experiments is used as the basis for the calculation models. The calculated results from SCALE 4.4a using the 27-group ENDF/B-IV, 44-group ENDF/B-V, and 238-group ENDF/B-V SCALE 4.4a cross-section libraries have been compared with reported results for the benchmark experiments. Comparison of results demonstrates that SCALE 4.4a run on the Nuclear Criticality Safety Division (NCSD) Hewlett Packard Series 9000/J-5600 (HP J-5600) unclassified workstation (referred to as CMODB) produces the same results within the statistical uncertainty of the Monte Carlo calculations as reported by the experiments.

Y/DD-896 provides assurance that the code predicts critical systems reasonably well, however, in general, there is wide variation in performance depending on the type of system and cross section library, based on bias results. Using a rating scale similar to Reference 2, the performance ranges from "very good" or less than about +/- 0.01 to "very poor" or about +0.035/-0.025 depending on the category of experiments. It is noted, however, that the best performance is somewhat subject to question because it is for the categories with relatively few experiments. Furthermore, areas of even worse performance are reported in the OECD handbook for some experiments, which were rejected due to known cross section problems (e.g. HEU-MET-FAST-035). In general, the observed problems resulted in very high bias, which would be conservative for plant applications. Nevertheless, there is no guarantee against the possibility of an unusual scenario that could produce non-conservative results, and analysts must be always vigilant for this possibility.

Y/DD-972 addresses various methods for determination of a USL and for incorporating uncertainty and margin into this USL. It provides comparisons of these methods and recommends a preferred method for determining the USL. Y/DD-972 establishes subcritical limits determined through an evaluation of statistics parameters of calculation results for critical experiments. However, the correlating parameters (i.e., mass, enrichment, geometry, absorption, moderation,

reflection, etc.) and values for applying additional margin to the subcritical limits are application dependent. The determination of correlating parameters and additional margin is an integral part of the process analysis. For the critical experiment results, no correlation between calculation results and neutron energy causing fission (or any other parameter) was found. As such, this document does not specify "final" USL values as has been done in the past.

The "Area of Applicability" (AoA) is also application dependent. In using the "raw" USL data of this report, the criticality analyst must exercise sound engineering judgment and use strong technical arguments to develop "a margin in k_{eff} or other correlating parameter that is sufficiently large to ensure that conditions (calculated by this method to be subcritical by this margin) will actually be subcritical."[1] Preparation of the "written report of validation" as identified in Section 4.3.6 of ANSI/ANS-8.1-1998 is the responsibility of the analyst.

4   Recommendations for Future Work

As stated in Y/DD-896, numerous additional OECD experiment descriptions are already available for experiments that would be directly applicable to Y-12 applications. In this regard, future efforts should be dedicated to the following: (1) developing more input decks in-house for experiments selected based on applicability to Y-12, and (2) including additional models as they become available in future revisions of the OECD report. In addition, changes in experiment categorization and interpretation/clarification sometimes are reported in the annual revisions of the OECD handbook. These changes should be reviewed to determine whether they impact the conclusions in this report. This work will require a dedicated effort.

Additional work may be needed to identify trends and/or biases by including additional parameters in the statistical evaluation of critical experiment results.

Additional work is needed on Area of Applicability (AoA) determination methods. Determining whether a particular application is within the AoA of a validation is currently almost exclusively based on subjective judgment. To be done correctly, it requires the analyst to have close familiarity with the details of the experiments in the validation. Efforts to develop input models for more experiments, particularly if done by several staff members, would have the additional benefit of increasing staff's familiarity with experiment descriptions, which would also help improve these AoA determinations. Doing such model input preparation work as part of the qualification program is suggested. In addition, methods have been proposed for more rigorous, and quantitative, AoA determinations. A dedicated effort is needed to evaluate these methods to determine if they are effective and practical to implement.

As discussed in Y/DD-896, based on the bias results, progressively better agreement between calculated k-effective results and expected results are obtained with cross section libraries having

increased number of cross section groups. Furthermore, areas of relatively poor performance are seen, probably due to cross section data issues. Efforts to produce improved cross section libraries for SCALE should be supported.

CCG-308 (Appendix A) allows for comparing the mean calculated k_{eff} alone (without adding $2\acute{o}$) directly with the USL, when certain conditions are met. This would be a simpler method than currently used (and proceduralized). The required conditions are generally met. Changing to this approach should be considered further.

Appendix A - Determination of the Lower Tolerance Limit Through Statistical Methods (CCG-380)

Statistical Discussion of the Critical Experiment Data

The critical experiment data consisted of 503 cases from the OECD Handbook.[1] The data from the benchmark experiments involving uranium represents a sufficiently wide range of enrichments and chemical forms to cover many existing or presently planned activities for the Y-12 site. The benchmark experiments evaluated were both critical and subcritical experiments. For each case, a total of 500 generations were run with 2000 neutrons per generation. The first 100 generations were skipped for the determination of the k_cal, the final k-effective after 800,000 neutron histories. Simulation results for these critical experiments to be used for validation of CSAS 25 in SCALE4.4a code package were compiled into an OUTPUT data set.[2]

The data set (OUTPUT) for this evaluation consisted of the following parameters:
     1.) lifetime - average life span of a neutron;
     2.) generation time - average time between successive neutron generations;
     3.) nu bar - the average total number of neutron produced per fission;
     4.) k_cal- the final k-effective after 800,000 neutron histories;
     5.) k_exp - the target k-effective for the experiment as modeled;
     6.) AFG - the average energy group at which fission occurs;
     7.) EALCF - the energy of the average lethargy of neutron causing fission;
     8.) k vs run - the average k-effective corresponding to the smallest mean standard deviation when the average k-effective and the mean standard deviation of the k-effective are computed for each generation over the range of nskip (user input number of generations to skip) through the total number of generations; and,
     9.) k vs skip - the average k-effective corresponding to the smallest mean standard deviation when the average k-effective and the mean standard deviation of the average k-effective are computed for the number of generations skipped over the range of nskip+1 through 2/3 the total number of generations calculated.

The following parts of the data set (OUTPUT) were used in the statistical evaluation: SCALE simulations identified by categorical variables of Number of Group Cross Sections (27, 44, and 238); Uranium-Enrichment (High or Low); and Physical Form (Compound, Metal, or Solution). The final k-effective (k_cal) results for the 27-group cross sections delineated by enrichment and physical form can be found in Table 1, the final k-effective (k_cal) results for the 44-group cross sections delineated by enrichment and physical form can be found in Table 2, and the final k-effective (k_cal) results for the 238-group cross sections delineated by enrichment and physical form can be found in Table 3.

Due to time constraints, dependent variables such as enrichment or mass used as INPUT to the simulation cases could not be quantified in the OUTPUT data set provided. Therefore, the final k-effective result (k_cal) could not be regressed on any of the input dependent variables such as

the amount of material or enrichment. In addition, the correlation of k-effective with AFG is $r^2 = 0.04$ for the 27-group cross sections, a $r^2 = 0.01$ for the 44-group cross sections, and $r^2 = 0.01$ for the 238-group cross sections. Similarly, the correlation of k-effective with EALCF is $r^2 = 0.02$ for the 27-group cross sections, $r^2 = 0.01$ for the 44-group cross sections, and $r^2 = 0.0003$ for the 238-group cross sections. These results are an indication that there is no linear relationship between the two variables. For this reason, regression and weighted regression do not seem to be applicable because of the lack of quantified independent variables in the data set (OUTPUT) and the low correlation with AFG and EALCF as discussed.

The distribution of the k_cal estimates is examined by the Wilk-Shapiro test[3] using the W-statistic. The W-statistic for a normal distribution ranges from 0 to 1 with small values leading to rejection of normality. The probability of a small W-statistic is given by Pr<W in Tables 1, 2, and 3. For Pr<W less than 0.05, we reject the null hypothesis that the k_cal estimates have a normal distribution. Otherwise, the k-cal estimates are assume to be normally distributed until additional data is collected to disprove the null hypothesis.

Tables 1, 2, and 3 list both a between standard deviation and a within standard deviation for the k_cal estimates. The between standard deviation represents the variability of the average k_cal estimates between the different criticality cases (e.g., for high enriched uranium metals, there are 118 cases). The between standard deviation of the k_cal is estimated by computing the variation between cases. This is performed by squaring each deviation in the k_cal, summing and dividing by "n-1", where "n" is the number of cases and taking the square root of the result. The within standard deviation represents the variability of the averaged k_cal value for a single simulation run. The within standard deviation is adjusted for the number of generations in the simulation and is estimated by pooling the within standard deviation in the different criticality cases. Thus, the within standard deviation of k_cal is the square root of the sum of the squares of the standard deviations of "n" cases divided by "n" number of cases

Table 1. The k_cal results for the 27- group cross sections.

Physical Form by Uranium Enrichment	Number of Simulations	Average k_cal	Between Standard Deviation	Within Standard Deviation	Minimum k_cal	Maximum k_cal	Probability of Normality Pr<W
High Enriched Uranium							
Metal	118	1.0048	0.0079	0.0010	0.9780	1.0349	0.0001
Compound	22	1.0000	0.0097	0.0010	0.9874	1.0185	0.1022
Solution	267	1.0063	0.0065	0.0012	0.9844	1.0332	0.6611
Combined Category for High Enriched Uranium	407	1.0055	0.0073	0.0011	0.9780	1.0349	0.0031
Low Enriched Uranium							
Metal	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Compound	87	0.9877	0.0056	0.0009	0.9769	0.9970	0.0001
Solution	9	0.9967	0.0021	0.0008	0.9934	1.0005	0.9933
Combined Category for Low Enriched Uranium	96	0.9886	0.0060	0.0009	0.9769	1.0005	0.0001
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	1.0023	0.0097	0.0011	0.9769	1.0349	0.0001

Table 2. The k_cal results for the 44-group cross sections.

Physical Form by Uranium Enrichment	Number of Simulations	Average k_cal	Between Standard Deviation	Within Standard Deviation	Minimum k_cal	Maximum k_cal	Probability of Normality Pr<W
High Enriched Uranium							
Metal	118	0.9987	0.0080	0.0010	0.9717	1.0307	0.0008
Compound	22	1.0015	0.0050	0.0010	0.9903	1.0101	0.6621
Solution	267	1.0039	0.0068	0.0011	0.9812	1.0310	0.0113
Combined Category for High Enriched Uranium	407	1.0023	0.0074	0.0011	0.9717	1.0310	0.0006
Low Enriched Uranium							
Metal	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Compound	87	0.9908	0.0054	0.0009	0.9804	0.9986	0.0001
Solution	9	0.9977	0.0021	0.0009	0.9944	1.0013	0.9736
Combined Category for Low Enriched Uranium	96	0.9915	0.0055	0.0009	0.9804	1.0013	0.0001
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	1.0002	0.0082	0.0011	0.9717	1.0310	0.0353

Table 3. The k_cal results for the 238-group cross sections.

Physical Form by Uranium Enrichment	Number of Simulations	Average k_cal	Between Standard Deviation	Within Standard Deviation	Minimum k_cal	Maximum k_cal	Probability of Normality Pr<W
High Enriched Uranium							
Metal	118	0.9974	0.0079	0.0010	0.9681	1.0284	0.0001
Compound	22	0.9978	0.0056	0.0010	0.9858	1.0085	0.7869
Solution	267	1.0031	0.0072	0.0012	0.9802	1.0301	0.0002
Combined Category for High Enriched Uranium	407	1.0011	0.0078	0.0011	0.9681	1.0301	0.0001
Low Enriched Uranium							
Metal	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Compound	87	0.9881	0.0060	0.0009	0.9751	0.9966	0.0001
Solution	9	0.9981	0.0025	0.0008	0.9941	1.0029	0.7133
Combined Category for Low Enriched Uranium	96	0.9890	0.0065	0.0009	0.9751	1.0029	0.0001
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	0.9988	0.0089	0.0011	0.9681	1.0301	0.0011

Tables 1, 2, and 3 partition the simulation runs by group cross sections, uranium enrichment and physical form. Additional partitions can also be made using the energy spectrum and cases. These additional partitions would reduce the between standard deviation of k_cal results but only a few or no cases would be available for each combination. Thus, additional partitions of the data set were not done.

II. Discussion of Benchmark Data

The benchmark data is provided in the OECD Handbook for each of the critical experiments evaluated. The benchmark (k_exp) or target k-effective is the best estimate of k-effective to be expected if the code performed a perfect calculation. Also, included in the OUTPUT data set is the uncertainty attached to the k_exp. Note, for the majority of critical experiments the k_exp is equal to one (1). Table 4 gives the statistical summary for the k_exp values.

Table 4. The k_exp values statistical summary.

Physical Form by Uranium Enrichment	Number of Simulations	Average k_exp	Between Standard Deviation	Minimum k_exp	Maximum k_exp
High Enriched Uranium					
Metal	118	0.9991	0.0027	0.9826	1.0028
Compound	22	0.9999	0.0022	0.9953	1.0060
Solution	267	0.9995	0.0014	0.9942	1.0015
Combined Category for High Enriched Uranium	407	0.9994	0.0019	0.9826	1.0060
Low Enriched Uranium					
Metal	N/A	N/A	N/A	N/A	N/A
Compound	87	0.9999	0.0001	0.9997	1.0000
Solution	9	1.0005	0.0002	1.0003	1.0009
Combined Category for Low Enriched Uranium	96	1.0000	0.0002	0.9997	1.0009
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	0.9995	0.0017	0.9826	1.0060

III.   Discussion of Bias Data

The bias estimated from the OUTPUT data set is defined as

$$\text{Bias} = k\_cal - k\_exp.$$

Table 5 gives the summary results for the bias results for the 27-group cross sections; Table 6 gives the summary results for the 44-group cross sections; and Table 7 gives the summary results for the 238-group cross sections. The signed-rank test was used to test the null hypothesis that the expected value of the bias is 0 (i.e., E[k_cal - k_exp] = 0) for a significance level of á = 0.05. The use of the signed-rank test for significance does not require that the populations be symmetric.[4] The signed-rank test for the bias results showed that for all partitions of the 27-group cross sections, except the high enriched compound result, are significantly different from zero. The sign-rank test for the bias results showed that for three partitions of the 44-group cross sections; highly enriched solution results, low enriched compound and solution results, are significantly different from zero. Likewise, the sign-rank test for the bias results showed that for all partitions of the 238-group cross sections except for the high enriched compound results were significantly different than zero. A positive average bias represents a high bias and a negative average bias represents a low bias.

Table 5. The bias results for the 27-group cross sections.

Physical Form by Uranium Enrichment	Number of Simulations	Average Bias	Standard Deviation Bias	Minimum Bias	Maximum Bias	Probability of Bias Test Signed-Rank Test
High Enriched Uranium						
Metal	118	0.0057	0.0067	-0.0092	0.0349	0.0001
Compound	22	0.0000	0.0085	-0.0114	0.0148	0.9126
Solution	267	0.0067	0.0065	-0.0156	0.0332	0.0001
Combined Category for High Enriched Uranium	407	0.0061	0.0069	-0.0156	0.0349	0.0001
Low Enriched Uranium						
Metal	N/A	N/A	N/A	N/A	N/A	N/A
Compound	87	-0.0122	0.0056	-0.0231	-0.0027	0.0001
Solution	9	-0.0038	0.0022	-0.0071	-0.0001	0.0039
Combined Category for Low Enriched Uranium	96	-0.0114	0.0059	-0.0231	-0.0001	0.0001
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	0.0027	0.0096	-0.0231	0.0349	0.0001

Table 6. The bias results for the 44-group cross sections.

Physical Form by Uranium Enrichment	Number of Simulations	Average Bias	Standard Deviation Bias	Minimum Bias	Maximum Bias	Probability of Bias Test Signed-Rank Test
High Enriched Uranium						
Metal	118	-0.0004	0.0067	-0.0154	0.0307	0.1520
Compound	22	0.0015	0.0042	-0.0097	0.0087	0.0650
Solution	267	0.0044	0.0070	-0.0188	0.0310	0.0001
Combined Category for High Enriched Uranium	407	0.0029	0.0071	-0.0188	0.0310	0.0001
Low Enriched Uranium						
Metal	N/A	N/A	N/A	N/A	N/A	N/A
Compound	87	-0.0091	0.0054	-0.0196	-0.0011	0.0001
Solution	9	-0.0029	0.0021	-0.0061	0.0008	0.0117
Combined Category for Low Enriched Uranium	96	-0.0085	0.0055	-0.0196	0.0008	0.0001
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	0.0007	0.0082	-0.0196	0.0310	0.0950

Table 7. The bias results for the 238-group cross sections.

Physical Form by Uranium Enrichment	Number of Simulations	Average Bias	Standard Deviation Bias	Minimum Bias	Maximum Bias	Probability of Bias Test Signed-Rank Test
High Enriched Uranium						
Metal	118	-0.0017	0.0064	-0.0203	0.0284	0.0003
Compound	22	-0.0021	0.0047	-0.0142	0.0073	0.0509
Solution	267	0.0035	0.0074	-0.0198	0.0301	0.0001
Combined Category for high Enriched Uranium	407	0.0017	0.0074	-0.0203	0.0301	0.0001
Low Enriched Uranium						
Metal	N/A	N/A	N/A	N/A	N/A	N/A
Compound	87	-0.0118	0.0060	-0.0247	-0.0031	0.0001
Solution	9	-0.0025	0.0025	-0.0064	0.0024	0.0273
Combined Category for Low Enriched Uranium	96	-0.0110	0.0088	-0.0247	0.0024	0.0001
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	-0.0007	0.0088	-0.0247	0.0301	0.0001

IV.   Methodology for the Determination of Tolerance Limits

Sometimes, one is more interested in the range of values in a population than in the average value. Statistical tolerance limits furnish limits between, above, or below which we confidently expect to find a prescribed proportion P of individual items of a population For example for the high enriched metal results, a lower tolerance limit statement would be "there is a probability of 0.95 that 95% of the population is greater than a specified result" for high enriched uranium metal results in the 27-group cross sections. Tolerance limits are not the same as confidence intervals which are probability statements about the parameters of a population (e.g., mean, variance, etc.). For our application, a new simulation result will be compared to a lower tolerance limit on a

k_cal for one of the partitions for the simulation cases. Note, future simulation k_cal results may fall below the lower tolerance limit. This situation is not an unreasonable because by definition at least 5% of the critical experiments have probability of lying below the lower tolerance limit.

Lower tolerance limits are usually calculated by two methods depending on the assumption of the probability distribution of the k_cal values. If the values are assumed to have a normal distribution, the tolerance interval has the form Average - A × (Standard Deviation) where A is a constant for the number of simulation cases, probability level, and population coverage. These A values can be found in M.G.Natrella, NBS Handbook 91 titled *Experimental Statistics* (Table A-7).[5] For the situation when a normal distribution is not appropriate, a non-parametric method is suggested. The non-parametric method gives the rank of the lowest data value that can be used for the population values. These rankings also are given in M.G.Natrella, NBS Handbook 91 titled *Experimental Statistics* (Table A-31).[5] A normal tolerance interval may not be as conservative because the interval is not as large as a non-parametric tolerance interval. The non-parametric method is appropriate for a wide range of probability distributions. An advantage of the normal tolerance interval is that it can be calculated for small sample sizes when the non-parametric method is not applicable.

In estimating the tolerance limits for the critical experiments, it was determined there were two options available. Option 1 represent simulations to duplicate criticality experiments. For this option, the k_sim results account for bias and the uncertainty for model errors and for new simulation results. The k_sim value for a simulation should be greater than a lower tolerance limit on the criticality experiment previously simulated. If the k_sim value is above the lower tolerance limit for the partition and group cross sections in Tables 9, 10, and 11, then the analyst can state with specified confidence that the new simulation can be considered part of the original group of criticality experiments. Reference Section VIII for an examples that illustrate the use of Option 1.

Option 2 represents simulations to find the k_pla for a plant scenario. For this option, the k_pla results account for bias and the uncertainty due to model errors, to new simulation results and to the selection of correct safety margin. The k_pla value for a plant simulation should be less the appropriate lower tolerance limit as specified in Tables 12, 13, and 14 assuming the partition correction factor (PCF) is 0.02. Reference Section VI.A for a discussion of the partition correction factor and Section VIII for examples that illustrate the use of Option 2.

A. Determination of One Sided Lower Tolerance Limits for Normal Distributions

A normal lower interval is of the form Average- A x (standard deviation). The average and the standard deviation are given in the tables in Section I. The A values can be found in M.G Natrella, NBS Handbook 91.[5] The A value is selected so that we have a level of confidence (e.g., 0.95, 0.99) that a proportion P of the k-effective population is above the lower tolerance limit.

B. Tolerance Limits Independent of the Form of the Distribution

If the k-effective results are non-normal, then assuming normality will have the effect that the true proportion P of the population above the lower tolerance limits will vary from the intended P. If we do not rely on the normal assumption, non-parametric lower tolerance limits can be calculated that only assume the distribution has no discontinuities. The lower tolerance limits so obtained may be substantially wider than those assuming normality.

Table 8 gives the number of samples required to calculate non-parametric lower tolerance limits for different combinations of confidence levels (0.95 and 0.99) and population proportion P (95% and 99%). For example, a lower confidence level of 0.95 on 95% of the population requires 59 data values and the lower tolerance limit would be the smallest data value. As the sample size increases, the r-th ordered data value (e.g., 2nd smallest, 3rd smallest) would be used for the lower tolerance limit depending on the number of data value. If the parameters are changed to a lower confidence level of 0.99 on 99% of the population, the number of samples required would be 459 and again the smallest data value would be the lower tolerance limit. Lower tolerance limits are given in Tables 9 to 14 for those partitions that have data with non-normal distributions. All of these partitions have sample sizes greater than the minimum value of 59 required for a 0.95 confidence interval on 95% of the population.

Table 8. Minimum number of samples required for lower non-parametric tolerance limits.

Population Proportion	Confidence Level	
	0.95	0.99
95%	59	90
99%	299	459

For lower non-parametric tolerance limits, Natrella [NBS-91 Table A-31][5] gives the smallest data value which is used for the lower tolerance limit for a given confidence limit and population proportion.

V.   Model Discussion for Option 1

A. Simulation Model

Simulations of the criticality safety benchmark experiments are run to determine if the KENO code results (k_cal) are in the population of previous simulated results. The

comparison of the new simulation result must take into account the uncertainty of the input parameters used to model a particular criticality safety benchmark experiment case as well as the simulation error.

The uncertainty for a new simulation run k_sim can be represented by the following model:

$$k\_sim = \text{True Value} + \text{Bias} + \text{Model Error} + \text{Simulation Error}$$

where

k_sim = The new simulated criticality safety benchmark experiment.

True Value = The true criticality value which is "1" for the criticality safety benchmark experiments. The majority of benchmark experiments were critical however some were subcritical and some were supercritical.

Bias = The difference "(k_cal - k_exp)" for k_cal previously simulated over a number of different models within a class of parameters from the known or expected k_exp. The model bias is assumed to be a constant which is specified by an average bias in Tables 5, 6, and 7.

Model Error = The simulated values are assumed to fall in one of the partition sets in Tables 1, 2, and 3 which are summary results over different model input parameter sets. The variation of the models within a partition set represents the uncertainty of the model error. The model error is assumed to have an expected value of zero and a variance of V_M. The model error standard deviation ($\sqrt{V_M}$) is estimated by the between standard deviation in Tables 1, 2, and 3.

Simulation Error = Simulation error is due to the variation within a simulation run. The simulation error has an expected value of zero and a variance V_s. The simulation error standard deviation ($\sqrt{V_s}$) is estimated by the within standard deviation in Tables 1, 2, and 3.

Expected Value

The expected value of the simulation run E[k_sim] is given as:

$$E[\,k\_sim\,] = 1 + \text{Bias}$$

This expected value is estimated by selecting the proper bias term from Tables 5, 6, or 7.

Variance Value

The variance value of the simulation run Var[k_sim] is given as:

$$\text{Var[k\_sim]} = V_M + V_S .$$

The variance term V_M is estimated from column 4 of Table 1, 2, or 3 by squaring the between error standard deviation. The variance term V_S is estimated from column 5 of Table 1, 2, or 3 by squaring the within error standard deviation.

Tolerance Interval

A lower tolerance interval is a lower limit for which a proportion "q" (e.g., 95%) of the population is greater than, with probability 1-á (e.g., 1 - 0.05 = 0.95). Tolerance intervals differ from confidence intervals because they are probability statements about population measurements while confidence intervals are probability statements about population parameters. A new simulated value can be judge to be a reasonable value if it is above the lower tolerance interval of the criticality safety benchmark experiments. If the new simulation value k_sim is above the lower tolerance limit on Table 9, 10, or 11, then the new simulation value can be considered part of the original sampled population.

Tolerance intervals can be calculated by two methods (a) assuming the simulation results have a normal distribution, and (b) non-parametric methods.

1. Normal Assumption:

$$Lower\ Tolerance\ Limit = (\ 1 + Average\ Bias\ ) - A\ x\ \sqrt{V_M + V_S}$$

where the factor A is found in M.G.Natrella, NBS Handbook 91 titled *Experimental Statistics* (Table A-7).[5]

2. Non-Parametric:

Find the r-th ordered data value for which a "P" population proportion is greater than, at the 1-á confidence level.

Order statistics can be found in M.G.Natrella, NBS Handbook 91 titled *Experimental Statistics*

(Table A-31).[5]

B. One Sided Lower Tolerance Limits

Table 9 gives the results of the tolerance bound calculations using univariate normal statistics and non parametric statistics as appropriate for the 27-group cross sections for the set of $k\_sim = 1 + bias$ results.

Table 10 gives the results of the 44-group cross sections using univariate normal statistics and non parametric statistics as appropriate for the 44-group cross sections for the set of $k\_sim = 1 + bias$ results.

Table 11 gives the results of the tolerance bound calculations using univariate normal statistics and non parametric statistics as appropriate for the 238-group cross sections for the set of $k\_sim = 1 + bias$ results.

The Nuclear Criticality Safety Division at the Y-12 site selected the lower non-parametric tolerance limit to be the smallest value of the data set. This selection guarantees that the lower tolerance limit will have a 0.95 lower confidence limit on 95% of the population because all non-normal partitions have sample sizes larger than 59 (see Table 8).

Table 9. The lower tolerance limit for the set of k_sim = 1 + bias results using normal and non-parametric distributions for the 27-group cross sections. Calculations for normal lower tolerance limits are for cases with a normal distribution (see Table 1).

Physical Form by Uranium Enrichment	Number of Simulations	0.95/95% Normal Lower Tolerance Limit	Minimum Value for Non-Parametric Lower Tolerance Limit (at least 0.95/95%)	Code Bias: 1 - Lower Tolerance Limit
High Enriched Uranium				
Metal	118	N/A	0.9908	0.0092
Compound	22	0.9771	N/A	0.0229
Solution	267	0.9947	N/A	-0.0015
Combined Category for High Enriched Uranium	407	N/A	0.9844	0.0156
Low Enriched Uranium				
Metal	N/A	N/A	N/A	N/A
Compound	87	N/A	0.9769	0.0231
Solution	9	0.9894	N/A	0.0106
Combined Category for Low Enriched Uranium	96	N/A	0.9769	0.0231
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	N/A	0.9769	0.0231

Table 10. The lower tolerance limit for the set of k_sim = 1 + bias results using normal and non-parametric distributions for the 44 group. Calculations for normal lower tolerance limits are for cases with a normal distribution (see Table 2).

Physical Form by Uranium Enrichment	Number of Simulations	0.95/95% Normal Lower Tolerance Limit	Minimum Value for Non-Parametric Lower Tolerance Limit (at least 0.95/95%)	Code Bias: 1 - Lower Tolerance Bound
High Enriched Uranium				
Metal	118	N/A	0.9846	0.0154
Compound	22	0.9895	N/A	0.0105
Solution	267	N/A	0.9812	0.0188
Combined Category for High Enriched Uranium	407	N/A	0.9812	0.0188
Low Enriched Uranium				
Metal	N/A	N/A	N/A	N/A
Compound	87	N/A	0.9804	0.0196
Solution	9	0.9902	N/A	0.0098
Combined Category for Low Enriched Uranium	96	N/A	0.9804	0.0196
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	N/A	0.9804	0.0196

Table 11. The lower tolerance limit for the set of k_sim = 1 + bias results using normal and non-parametric distributions for the 238 cross sections group. Calculations for normal lower tolerance limits are for cases with a normal distribution (see Table 3).

Physical Form by Uranium Enrichment	Number of Simulations	0.95/95% Normal Lower Tolerance Limit	Minimum Value for Non-Parametric Lower Tolerance Limit (at least 0.95/95%)	Code Bias: 1 - Lower Tolerance Bound
High Enriched Uranium				
Metal	118	N/A	0.9797	0.0203
Compound	22	0.9845	N/A	0.0155
Solution	267	N/A	0.9802	0.0198
Combined Category for High Enriched Uranium	407	N/A	0.9797	0.0203
Low Enriched Uranium				
Metal	N/A	N/A	N/A	N/A
Compound	87	N/A	0.9753	0.0247
Solution	9	0.9894	N/A	0.0106
Combined Category for Low Enriched Uranium	96	N/A	0.9753	0.0247
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	N/A	0.9753	0.0247

VI. Model Discussion for Option 2

A. Simulation Model

The KENO code can be used to simulate actual plant scenarios. The new simulation result from an actual plant scenario, k_pla must take into account the uncertainty of the input parameters used for the model as well as the simulation error of the model. The

uncertainty for a new simulation run for an actual plant scenario can be represented by the
following model:

k_pla = True Value + Bias + Model Error + Simulation Error + Partition Error

where

k_pla	=	The simulated plant scenario.

True Value   =   The true criticality value which is "1" if the simulated plant scenario
is critical or less than 1 if the plant scenario is sub-critical.

Bias   =   The difference "(k_cal - k_exp)" for k_cal previously simulated
over a number of different models within a class of parameters from
the known or expected k_exp. The model bias is assumed to be a
constant which is specified by an average bias in Tables 5, 6, and 7.

Model Error   =   The simulated values are assumed to fall in one of the partition sets
in Tables 1, 2, and 3 which are summary results over different
model input parameter sets. The variation of the models within a
partition set represents the uncertainty of the model error. The
model error is assumed to have an expected value of zero and a
variance of V_M. The model error standard deviation ($\sqrt{V_M}$) is
estimated by the between standard deviation in Tables 1, 2, and 3.

Simulation Error =   Simulation error due to the variation within a simulation run. The
simulation error is assumed to have an expected value of zero and a
variance of V_S. The simulation error standard deviation ($\sqrt{V_s}$) is
estimated by the within standard deviation in Tables 1, 2, and 3.

Partition Error =   The partition error accounts for the difference between the plant
scenario and the selected partition used for comparison. The plant
simulation may be very different then the selected partition used for
comparison. The variation of the partition error can not be
estimated but is reflected in the lower tolerance interval by
subtracting a partition correction factor (PCF). Typically, the PCF
ranges from 0.02 to 0.05 and is selected by the analyst based on
engineering judgement.

Expected Value

The expected value of the simulation run E[k_pla] is given as:

$$E[k\_pla]= k\_pla + Bias$$

This expected value is estimated by selecting the proper bias term from Table 5, 6, or 7, and the k_pla term is specified in the output from the KENO code for the plant scenario.

Variance Value

The variance value of the plant simulation run Var[k_pla] is given as:

.
$$Var[k\_pla] = V_M + V_S.$$

The variance term V_M is estimated from column 4 of Table 1, 2, or 3 by squaring the between error standard deviation. The variance term V_S is estimated from column 5 of Table 1, 2, or 3 by squaring the within error standard deviation.

Tolerance Interval

A lower tolerance interval is a lower limit for which a proportion P (e.g., 95%) of the population is greater than, with probability 1-á (e.g., 1 - 0.05 = 0.95). A new plant simulated value can be judge to be a subcritical value if it is below the lower tolerance limit determination as shown in Tables 9, 10, and 11. In addition, the lower tolerance limit is adjusted by subtracting the partition correction factor. The PCF is a number chosen by the analyst that typically ranges from 0.02 to 0.05. The partition error accounts for the difference between the plant scenario and the selected partition used for comparison.

$$Lower\ Tolerance\ Limit = (\ 1 + Bias\ ) - A\ x\ \sqrt{V_M + V_S} - PCF$$

1.   Normal Assumption:

where

 the factor A is found in M.G.Natrella, NBS Handbook 91 titled *Experimental Statistics* (Table A-7).[5]

2.   Non-parametric:

Find the r-th ordered measurement for which a P proportion of the population is greater than, with 1-á probability as discussed in Section V.A. In addition, the partition error must be accounted for by the subtracting the PCF from the lower tolerance limit. The partition error accounts for the difference between the plant scenario and the selected partition used for comparison.

VII.   Recommendations for Future Validation Statistical Evaluations

A nested design of the categorical variables gives indication of possible trends in the data. Thus, an experimental design based on the INPUT parameters of the Scale Code should be designed to analyze the most significant effects on the simulation values. These designs can be fractional factorial, Plackett-Burman, or highly saturated designs to examine the largest number of INPUT parameters in the least number of simulation runs.

VIII.   Examples

A. Option 1

For the following examples the assumptions are as follows:

1)   The PCF is not required;
2)   An adequate number of generations and neutrons per generation have been run for the experiment selected; and
3)   The simulation converges for the experimental parameter results.

Example 1

Assume the analyst would like to include the following critical experiment HEU-MET-FAST-035 in the group (k_cal) already evaluated. The k_sim (the new simulated k_cal) result for this experiment run on the work station with the 238 cross section group was 1.0915. The k_exp result as documented in the OECD is 0.9982.

The analyst would check Table 11 for the High Enriched Uranium Metal partition. The lower tolerance limit is 0.9797. Because 1.0915 is greater than 0.9797 the analyst could assume this experiment is part of the group evaluated. Note, the simulation result (i.e 1.0915) is larger than the maximum critical experiment result (i.e. 1.0284, see Table 3) for this group, inclusion of this experiment may not be valid. In this case, determination of a two sided tolerance limit is recommended

Example 2

Assume the analyst would like to include the following critical experiment HEU-MET-FAST-055 in the group (k_cal) already evaluated. The k_sim result for this experiment run on the work station with the 238 cross section group was 1.0063. The k_exp result as documented in the OECD is 0.9982.

The analyst would check Table 11 for the High Enriched Uranium metal partition system . The lower tolerance limit is 0.9797. Because 1.0063 is greater than 0.9797 the analyst could assume this experiment is part of the group evaluated. In this instance because the result (i.e. 1.0063) is well below the greatest value for that partition (i.e. 1.0284, see Table 3), the assumption that this critical experiment is part of the set originally evaluated is valid.

B. Option 2

For the following examples the assumptions are as follows:

1)   The PCF is required;
 2) An adequate number of generations and neutrons per generation have been run for the experiment selected; and
3) The simulation converges for the experimental parameter results; and
4) The within standard deviation is less than the between standard deviation of the partition selected.

Example 1

Suppose an analyst ran a high enriched uranium solution case with 238 cross section group with the result k_pla = 0.9701 which was simulated with only 50 generations.

A check of Table 14 shows that the lower tolerance limit for the 238 group cross section assuming PCF = 0.02 is 0.9602 which implies this experiment may be in the population of critical experiments.

Example 2

Suppose an analyst ran a high enriched uranium solution case with 238 cross section group with the result k_pla = 0.9601 with a within standard deviation of 0.0009.

A check of Table 14 shows that the lower tolerance limit for the 238 cross section group assuming PCF = 0.02 is 0.9602 which is greater than 0.9601. The lower tolerance limit depends on the variation of the data quantified by the between standard deviation. The larger the between standard deviation the smaller the lowest data value and hence the lower tolerance limit. The analyst should compare the simulated standard deviation to see if it is larger than the between standard deviation to be reassured that the lower tolerance limit are not based on data with an under estimated variation. For this example, the plant scenario with a simulation standard deviation of 0.0009 is less than the between standard deviation of 0.0072. Therefore, the analyst could assume the plant scenario is not a member of the partition selected or is sub-critical.

Table 12. The lower tolerance limit for the set of k_sim = 1 + bias results using normal and non-parametric distributions for the 27 cross sections group. Calculations for normal lower tolerance limits are for cases with a normal distribution (see Table 1).

Physical Form by Uranium Enrichment	Number of Simulations	0.95/95% Normal Lower Tolerance Limit	Minimum Value for Non-Parametric Lower Tolerance Limit (at least 0.95/95%)	Subtracted PCF = 0.02
High Enriched Uranium				
Metal	118	N/A	0.9908	0.9708
Compound	22	0.9771	N/A	0.9571
Solution	267	0.9947	N/A	0.9747
Combined Category for High Enriched Uranium	407	N/A	0.9844	0.9644
Low Enriched Uranium				
Metal	N/A	N/A	N/A	N/A
Compound	87	N/A	0.9769	0.9569
Solution	9	0.9894	N/A	0.9694
Combined Category for Low Enriched Uranium	96	N/A	0.9769	0.9569
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	N/A	0.9769	0.9569

Table 13. The lower tolerance limit for the set of k_sim = 1 + bias results using normal and non-parametric distributions for the 44 cross sections group. Calculations for normal lower tolerance limits are for cases with a normal distribution (see Table 2).

Physical Form by Uranium Enrichment	Number of Simulations	0.95/95% Normal Lower Tolerance Limit	Minimum Value for Non-Parametric Lower Tolerance Limit (at least 0.95/95%)	Subtracted PCF = 0.02
High Enriched Uranium				
Metal	118	N/A	0.9846	0.9646
Compound	22	0.9895	N/A	0.9695
Solution	267	N/A	0.9812	0.9612
Combined Category for High Enriched Uranium	407	N/A	0.9812	0.9612
Low Enriched Uranium				
Metal	N/A	N/A	N/A	N/A
Compound	87	N/A	0.9804	0.9604
Solution	9	0.9902	N/A	0.9702
Combined Category for Low Enriched Uranium	96	N/A	0.9804	0.9604
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	N/A	0.9804	0.9604

Table 14. The lower tolerance limit for the set of k_sim = 1 + bias results using normal and non-parametric distributions for the 238 cross sections group. Calculations for normal lower tolerance limits are for cases with a normal distribution (see Table 3).

Physical Form by Uranium Enrichment	Number of Simulations	0.95/95% Normal Lower Tolerance Limit	Minimum Value for Non-Parametric Lower Tolerance Limit (at least 0.95/95%)	Subtracted PCF = 0.02
High Enriched Uranium				
Metal	118	N/A	0.9797	0.9597
Compound	22	0.9845	N/A	0.9645
Solution	267	N/A	0.9802	0.9602
Combined Category for High Enriched Uranium	407	N/A	0.9797	0.9597
Low Enriched Uranium				
Metal	N/A	N/A	N/A	N/A
Compound	87	N/A	0.9753	0.9553
Solution	9	0.9894	N/A	0.9694
Combined Category for Low Enriched Uranium	96	N/A	0.9753	0.9553
Combined Category for Both High Enriched Uranium and Low Enriched Uranium	503	N/A	0.9753	0.9553

IX.   References

1.   J.Blair Briggs, Editor, *International Handbook of Evaluated Criticality Safety Benchmark Experiments*, NEA/NSC/DOC(95)03, Volumes 1-7, Organization for Economic Cooperation and Development-Nuclear energy Agency, Paris, France, September 1999 Edition.

2.   R.H.Smith, J.F.DeClue, A.W. Krass, V.R.Cain, *Bias and Uncertainty of Critical Experiment Models with CSAS25 from SCALE4.4a for Criticality Safety Analyses On the HP J-5600 (CMODB) Workstation*, Y/DD-896, BWXT Y-12,L.L.C., Oak Ridge, Tennessee, January 2001.

3.   V.L.Anderson, R.A.Mclean, *Design of Experiments* (New York, Marcel Dekker, Inc., 1974).

4.   G.W.Snedecor, W.G.Cochran, *Statistical Methods* (Ames, Iowa, Iowa state University Press),1967.

5.   M.G.Natrella, *Experimental Statistics*, National Bureau of Standards Handbook 91, U.S. Government Printing Office, Washington, 1963.

5   REFERENCES

1 *Nuclear Criticality Safety in Operations with Fissionable Material Outside Reactors*, ANSI/ANS-8.1-1998, American National Standards Institute, New York, 1998.